

A Noisy Principal Component Analysis for Forward Rate Curves

Márcio Poletti Laurini^{a,*}, Alberto Ohashi^b

^a*FEA-RP USP - 14040-905, Ribeirão Preto, SP, Brasil*

^b*Mathematics Dept., Universidade Federal da Paraíba, 13560-970, João Pessoa, PB, Brasil*

Abstract

Principal Component Analysis (PCA) is the most common nonparametric method for estimating the volatility structure of Gaussian interest rate models. One major difficulty in the estimation of these models is the fact that forward rate curves are not directly observable from the market so that non-trivial observational errors arise in any statistical analysis. In this work, we point out that the classical PCA analysis is *not* suitable for estimating factors of forward rate curves due to the presence of measurement errors induced by market microstructure effects and numerical interpolation. Our analysis indicates that the PCA based on the long-run covariance matrix is capable to extract the true covariance structure of the forward rate curves in the presence of observational errors. Moreover, it provides a significant reduction in the pricing errors due to noisy data typically founded in forward rate curves.

Keywords: Finance, Pricing, Principal component analysis, term-structure of interest rates, HJM models.

1. Introduction

The term-structure of interest rates is a high-dimensional object which has been the subject of much research in the finance literature. It is the natural starting point for pricing fixed-income securities and other financial assets. In particular, the identification of factors capable to explain its movements plays a crucial role in modeling complex interest rate derivative products. Since the seminal works of Steeley (1990), Stambaugh (1988) and Litterman and Scheinkman (1991), it is well-known that most of the covariance yield curve structure can be summarized by just a few unobservable factors. This stylized fact is fundamentally based on the Principal Component Analysis (henceforth abbreviated by PCA) based on sample covariance matrices. In this case, a small number of eigenvectors summarizes the whole second moment structure of the yield curves.

^{*}We would like to thank Josef Teichmann for stimulating discussions on the topic of this paper. The second author was supported by CNPq grant 308742.

^{*}Tel.: +55-16-33290867

URL: mplaurini@gmail.com (Márcio Poletti Laurini), ohashi@mat.ufpb.br (Alberto Ohashi)

The interest rate markets can be summarized by two fundamental high dimensional objects: the yield $x \mapsto y_t(x)$ and forward rate curves $x \mapsto r_t(x)$; $t \geq 0$ which are connected by the following linear relation

$$y_t(x) = \frac{1}{x} \int_0^x r_t(z) dz; 0 \leq t < \infty, x \geq 0. \quad (1.1)$$

See e.g. Filipovic (2009) for more details. In particular, the underlying covariance structure of yield and forward rate curves play a major role in the statistical analysis of the term-structure of interest rate. See e.g. Rebonato (2002), Schmidt (2011) and other references therein. For instance, forward rate curves play a central role in pricing and hedging interest rate derivatives by means of the classical methodology proposed by Heath et al. (1992). Their contribution can be summarized by the representation of the forward rate curve dynamics in terms of a stochastic partial differential equation

$$dr_t(x) = \left(\frac{\partial r_t(x)}{\partial x} + \alpha_{HJM}(t, r_t(x)) \right) dt + \sum_{j=1}^d \sigma_j(t, r_t(x)) dB_t^j; r_0(x) = \xi(x), 0 \leq t < \infty, x \geq 0 \quad (1.2)$$

where α_{HJM} is the so-called (HJM) drift condition which is fully determined by the volatility structure $\sigma = (\sigma^1, \dots, \sigma^d)$, (B^1, \dots, B^d) is a d -dimensional Brownian motion and $x \mapsto \xi(x)$ is a given initial forward rate curve. Then, the initial forward rate curve ξ and the volatility structure σ fully determine the no-arbitrage dynamics of the model (1.2). See e.g. Filipovic (2009) for further details.

Plenty of spot interest rate data (and hence yield curve data) are available in fixed income markets. However, due to the absence of explicit forward rate markets, implied forward rate curves have to be estimated from interest rates based on other financial instruments. This already presents a major difficulty in implementing derivative pricing models based on the classical Heath-Jarrow-Morton methodology.

The most common non-parametric procedure for estimating the covariance structure of forward rate curves is the PCA methodology. Basically, three common strategies are very popular in the PCA estimation of the forward rate curves: **(A)** One postulates the existence of a finite-dimensional parameterized family of smooth curves $\mathcal{G} = \{G(z; x); z \in \mathcal{Z} \subset \mathbb{R}^N, x \geq 0\}$ and a \mathcal{Z} -valued state process Y such that

$$y_t(x) = G(Y_t; x) \text{ for } x \geq 0, 0 \leq t < \infty. \quad (1.3)$$

By interpolating the available yield data based on \mathcal{G} , then one extracts the associated forward rate curve by means of any numerical scheme to recover $r_t(x) = y_t(x) - x \frac{\partial y_t(x)}{\partial x}$. The PCA

is then applied on this estimated forward rate curves, as discussed in e.g. Jarrow (2002) and Lord and Pelsser (2007). **(B)** Instead of (1.3), one shall use a non-parametric polynomial splines method to interpolate the yield data and do the same step of **(A)**. See e.g. Vasicek and Fong (1982), Barzanti and Corradi (1998) and Chiu et al. (2008) for further details. Alternatively, one can use proxies to construct the forward rate curve jointly with a given interpolating family of smooth curves \mathcal{G} . See e.g. Bhar et al. (2002), Alexander and Lvov (2003) and Gauthier and Simonato (2012) for further details.

For a given initial forward rate curve, the fundamental object which encodes the whole dynamics of (1.2) is volatility. In particular, due to closed form expressions for derivative prices and hedging, it is common (see e.g. Rutkowski (1996), Jarrow (2002) and Falini (2010)) to assume that the volatility structure is deterministic. In this case, the stochastic dynamics of forward rates is given by a Gaussian HJM model:

$$dr_t(x) = \left(\frac{\partial r_t(x)}{\partial x} + \alpha_{HJM}(x) \right) dt + \sum_{j=1}^d \sigma_j(x) dB_t^j; \quad r_0(x) = \xi(x), \quad 0 \leq t < \infty, \quad x \geq 0. \quad (1.4)$$

The most common alternative to estimate the underlying volatility structure is to use PCA methodology (see e.g. Jarrow (2002) Filipovic (2009), Schmidt (2011) and other references therein) based on the static covariance matrix of a given sample $(r_t(x_1), \dots, r_t(x_n))$. The PCA methodology provides the following estimator for the volatility structure

$$\hat{\sigma}^i = \hat{\varphi}_i \sqrt{\hat{\lambda}_i}; \quad i = 1, \dots, \hat{m}, \quad (1.5)$$

where \hat{m} is the estimated number of principal components of the forward rate curves and the estimated eigenvalues and eigenvectors of the associated static covariance matrix are given by $\hat{\lambda}_i$ and $\hat{\varphi}_i$, respectively. One fundamental assumption behind (1.5) and, more generally, on the use of PCA in forward rate curves is the following one:

Assumption (I) There is no observational errors in forward rate curves.

At this stage, a natural question is the validity of assumption **(I)** in the term-structure of interest rates. In fact, we shall compare the existing literature of principal components between yield and forward rate curves to see some evidence of violation of assumption **(I)**. In one hand, the linearity of the relation (1.1) strongly suggests that dimension of the forward rate and yield curves must be identical (See Proposition 2.1). On the other hand, apparently, distinct results in the literature have been reported on the spectral structure of the forward rate and yield curves. Akahori et al. (2006) and Liu (2010) report a remarkable difference in the estimated number of factors between forward rate and yield curves and they suggest that a possible explanation for this would be the violation

of the random walk hypothesis. The same type of behavior was reported by Lekkos (2000) who argues that averaging the forward rates over time to maturities would induce a strong dependence on the yield data. He argues that PCA method artificially estimates a small number of principal components for yield curves. Alexander and Lvov (2003) study statistical properties of the UK Libor rates. They show that the strategy **(A)** induces equivalent loading factor structures between yield and implied forward rate curves. Lord and Pelsser (2007) report a visible difference in the PCA of forward and yield curves by using estimated Svensson curves for the Euro market.

Essentially, the existing literature restricts the discussion into two lines: (i) More factors are needed to account the correlation in forward rates curves. An averaging effect would be the reason for an artificial dependence on the yield curves. (ii) One way to remedy this pattern is to first smooth the yield data by means of a parametric or non-parametric form and then to calculate the implied forward rates. One should notice that one important assumption behind (i) and (ii) is **(I)**. More importantly, the current literature only suggests *ad hoc* methods based on **(A-B)**.

In this article, we take a rather different strategy. Throughout this article, we assume that the observed curve time series, which we denote by $X_1(\cdot), \dots, X_n(\cdot)$, they are subject to errors in the sense that

$$X_t(u) = r_t(u) + \varepsilon_t(u); \quad u \geq 0, t = 0, 1, \dots \quad (1.6)$$

The possible existence of the noise term ε in (1.6) reflects the classical interpolation procedures **(A-B)** when extracting forward rate curves from yield data. It can also be induced by observational errors due to market microstructure effects. The existence of an underlying bid and ask bond price structure contaminates the yield and forward rate curves (see e.g. Mizrahi and Neely (2011) and Goyenko et al. (2011)). These noisy discrete data are smoothed to provide “observed” curves $x \mapsto X_t(x)$ where both $r_t(\cdot)$ and $\varepsilon_t(\cdot)$ are unobservable. We investigate in detail the existence and the impact of observational errors $\{\varepsilon_t(x); t \in \mathbb{N}, x \in \mathbb{R}_+\}$ in (1.6) in the classical PCA methodology.

We show that market microstructure effects and common interpolation procedures **(A-B)** induce noisy forward rate curves which cause severe bias in the PCA methodology. The starting point of our analysis is the fact that the ranks of the covariance operators of the forward rate and yield curves are identical (see Proposition 2.1). In addition, we show that PCA based on the so-called *long-run covariance matrix* (henceforth abbreviated by LRCM) significantly improves the estimation of the covariance structure of forward rate curves. The impact of noisy data in pricing interest rate derivatives is also discussed.

This article is organized as follows. In Section 2, we report an elementary result about the equivalence of ranks between the covariance operators of the forward rate and yield curves. In Section 3, we describe some alternatives of estimating covariance structures in the presence of observational errors, the so-called LRCM estimators. Section 4 presents a detailed simulation

analysis reporting the performance of the PCA based on LRCM estimators, as well as, the role played by measurement errors in the PCA methodology applied to the term-structure of interest rate. In order to compare the simulation results with a real data set, Section 5 provides an empirical analysis on the number of principal components for US and UK term-structure on interest-rates. In Section 6, we analyse the impact of neglecting observational errors in pricing interest rate derivatives in light of the PCA methodology. Section 7 presents the final remarks.

2. Rank equivalence in covariance operators for forward rate and yield curves

In this section, we give a simple remark showing the number of principal components in forward rate curves should be exactly the same of the yield curves under some mild conditions. In spite of its simplicity, it is the starting point to investigate the violation of assumption **(I)** and it also gives a comparison criteria for our statistical analysis.

Let $\{P(t, T); (t, T) \in \Delta^2\}$ be the term-structure of bond prices where $\Delta^2 := \{(t, T); 0 \leq t \leq T < \infty\}$. Let

$$s(t, T) := \frac{-\log P(t, T)}{T - t}; (t, T) \in \Delta^2, \quad (2.1)$$

be the spot-interest rate prevailing at time t for maturity T and let $y_t(x) := s(t, t + x)$ be the correspondent yield curve at time indexed by the time to maturity $x = T - t$.

The forward rate prevailing at time t for maturity T is

$$f(t, T) := -\frac{\partial \log P(t, T)}{\partial T}; (t, T) \in \Delta^2,$$

and the forward rate curve is $r_t(x) := f(t, t + x)$ for $x = T - t$. For simplicity of exposition, we analyze the PCA methodology based on a space E of curves so that forward rate and yield data are interpreted as sample curves over time.

In the sequel, we consider a discrete-time setup $t \in \mathbb{N} := \{0, 1, 2, \dots\}$ and we assume that y_t and r_t are discrete-time weakly stationary E -valued process. That is, there exist functions $(\mu_y(\cdot), \mu_r(\cdot), Q_r(\cdot, \cdot), Q_y(\cdot, \cdot))$ such that the following identities hold for every t

$$\mu_r(u) = \mathbb{E}r_t(u), \quad \mu_y(u) = \mathbb{E}y_t(u),$$

$$Q_y(u, v) = \text{Cov}(y_t(u), y_t(v)), \quad Q_r(u, v) = \text{Cov}(r_t(u), r_t(v)); u, v \in K. \quad (2.2)$$

Otherwise, we assume that the first difference process satisfies such properties. The set E is a separable Hilbert space of functions from a bounded set $K := [0, x^*) \subset \mathbb{R}_+$ to \mathbb{R} such that $J, T : E \rightarrow E$ defined by

$$f \mapsto J_x(f) := \frac{1}{x} \int_0^x f(y) dy; \quad f \mapsto T_x(f) := x \frac{d}{dx} f(x), \quad (2.3)$$

are bounded linear operators. In the sequel, we denote E equipped with an inner product $\langle \cdot, \cdot \rangle^{1/2} = \|\cdot\|$. Assuming that the discrete-time processes r and y are square-integrable, the covariance operators induced by the kernels in (2.2) admit spectral decompositions over E as follows

$$Q_y(\cdot) = \sum_{k=1}^{\infty} \lambda_k(y) \langle \cdot, \varphi_k(y) \rangle \varphi_k(y), \quad Q_r(\cdot) = \sum_{k=1}^{\infty} \lambda_k(r) \langle \cdot, \varphi_k(r) \rangle \varphi_k(r),$$

where $(\varphi_k(r))_{k=1}^{\infty}$ and $(\varphi_k(y))_{k=1}^{\infty}$ are orthonormal bases for E with eigenvalues $(\lambda_k(r))_{k=1}^{\infty}$ and $(\lambda_k(y))_{k=1}^{\infty}$, respectively. The number of principal components of the forward and yield curves are the number of non-zero eigenvalues of Q_r and Q_y , respectively. We assume that $\lambda_1(r) > \dots > \lambda_p(r) > \lambda_{p+i} = 0$; $\lambda_1(y) > \dots > \lambda_q(y) > \lambda_{q+i} = 0$ for every $i \geq 1$, $\max(p, q) < \infty$ so that

$$\mathbb{E}\|r_t - \mu_r\|^2 = \sum_{i=1}^p \lambda_i(r), \quad \mathbb{E}\|y_t - \mu_y\|^2 = \sum_{i=1}^q \lambda_i(y); \quad t \in \mathbb{N}.$$

The explained variance associated with the k -th principal component for r and y can be expressed, respectively, by

$$\frac{\lambda_k(r)}{\sum_{i=1}^p \lambda_i(r)}, \quad \frac{\lambda_k(y)}{\sum_{i=1}^q \lambda_i(y)}.$$

In the sequel, we denote $\Lambda := I_d - T$ where I_d is the identity operator and T is given by (2.3). For a given bounded linear operator G , the correspondent self-adjoint operator will be denoted by G^* . The following simple remark shows how the covariance operators of the forward rate and yield curves are related to each other.

Proposition 2.1. *Let y and r be square-integrable weakly-stationary E -valued discrete-time processes. Assume that Q_r and Q_y are finite-rank operators and E is a Hilbert space realizing (2.3). Then $Q_r = \Lambda Q_y \Lambda^*$ and $Q_y = J Q_r J^*$. In particular, $\text{Rank } Q_r = \text{Rank } Q_y$.*

Proof: We recall that Q_r and Q_y are the unique self-adjoint, non-negative and bounded operators such that

$$\langle Q_r f, g \rangle = \mathbb{E} \langle r_t - \mu_r, f \rangle \langle r_t - \mu_r, g \rangle \quad (2.4)$$

$$\langle Q_y f, g \rangle = \mathbb{E} \langle y_t - \mu_y, f \rangle \langle y_t - \mu_y, g \rangle; \quad f, g \in E, t \in \mathbb{N}. \quad (2.5)$$

Moreover, we shall write $y_t(x) = J_x(r_t)$ and $r_t(x) = \Lambda_x(y_t)$. This fact, together with relations (2.4) and (2.5) and the continuity assumptions on Λ and J allow us to conclude that $Q_r = \Lambda Q_y \Lambda^*$ and $Q_y = J Q_r J^*$. Now let us define the following finite-rank, non-negative and self-adjoint operators

$$\bar{Q}_r := Q_y \Lambda^* \Lambda, \quad \bar{Q}_y := Q_r J^* J.$$

By the very definition, \bar{Q}_y and Q_y share the same non-zero eigenvalues and therefore,

$$\text{Rank } Q_y = \text{Rank } \bar{Q}_y \leq \text{Rank } Q_r J^* \leq \text{Rank } Q_r.$$

The same argument applies to \bar{Q}_r and Q_r so that $\text{Rank } Q_r \leq \text{Rank } Q_y$. This concludes the proof.

Despite the simplicity of Proposition 2.1, it provides an important information on dimension reducing techniques for forward rate and yield curves based on PCA: The effective number of principal components should be the same for forward rates and yield curves.

2.1. Implications to related work

Proposition 2.1 contradicts the heuristic argument given by Lekkos (2000) who argues that the identity $y = J(r)$ would smooth the spectral structure of the yield curve covariance operator. It also shed some light on the results reported by many authors who compare the PCA for forward and yield curves. In one hand, Akahori and Liu (2011), Liu (2010), Lord and Pelsser (2007) and Lekkos (2000) report substantial differences between the correlation structure of forward and yield curves. On the other hand, Alexander and Lvov (2003) report a very stable estimation by using discretely compounded forward rates and the parametric Svensson family to extract the forward rates. Their empirical results together with Proposition 2.1 might suggest that in order to extract the spectral properties of forward rates from the observed bond prices, a reasonable choice of a parametric family of smooth curves is a good starting point. However, one should notice that this procedure may potentially introduce an a priori loading factor structure on forward rates. As pointed out by Lord and Pelsser (2007), nonparametric procedures based on splines and bootstrap techniques might introduce a non-negligible noise on the forward curve.

Proposition 2.1 implies that a remarkable difference in the number of principal components between forward and yield curves is a strong evidence for the presence of unobserved noise in the data. If this is the case, then unobserved errors might induce a nontrivial bias in the statistical analysis of the forward rate curves. This remark will be the starting point for our analysis in the remainder of this article.

3. PCA based on LRCM and noisy data

One way to overcome the presence of observational errors in a PCA methodology is the so-called long-run covariance matrix (LRCM). We recall that the PCA method based on the usual static sample covariance matrix is only valid for independent and weakly stationary processes. In

the presence of some sort of temporal dependence caused by serial correlation or a contamination process, the use of the static sample covariance matrix is not the correct one anymore.

One alternative to correct those types of dependence is to use estimators based on the LRCM. In the sequel, we denote the process of interest as a vector w_t , and assume weak stationarity and ergodicity of w_t . The LRCM of w is defined by

$$V_{lr} := \lim_{n \rightarrow \infty} \text{var}(\sqrt{n}\bar{w}) = \sum_{j=-\infty}^{\infty} \gamma(j),$$

where $\gamma(j) := E[(w_t - E[w_t])(w_{t-j} - E[w_{t-j}])^\top]$ is the cross covariance in lag j , \bar{w} is the sample mean and \top denotes the transpose of a matrix. One important and standard case is $\gamma(j)$ at $j = 0$. In this case, all the non-contemporaneous cross-variances are equal to zero and we retrieve the usual static covariance matrix $V_s := \gamma(0)$ whose the usual estimator will be denoted by \hat{V}_s .

Although it is possible to estimate independently each covariance term $\gamma(j)$ by the correspondent sample quantities $\hat{\gamma}(j)$, the natural estimator of the long run matrix $\hat{V}^n := \sum_{j=-(n-1)}^{n-1} \hat{\gamma}(j)$ is not consistent because the number of parameters grows proportional to the sample size. In order to overcome this problem, a general non-parametric class of LRCM consistent estimators is introduced by Andrews (1991)

$$\hat{V}_{lr}^A := \sum_{j=-(n-1)}^{(n-1)} \alpha(j) \hat{\gamma}(j), \quad (3.1)$$

where $\alpha(j)$ is a sequence of weights of the form $\alpha(j) = K(j/b)$, where $K(\cdot)$ is a continuous symmetric kernel function such that $K(0) = 1$ and b a suitable bandwidth parameter such that $b \rightarrow \infty$ as $n \rightarrow \infty$.

Some optimal choices for the kernel function and the bandwidth parameter in (3.1) are discussed by Andrews (1991). The most important one is the bandwidth parameter. Precise conditions for consistency of the LRCM estimators of type (3.1) based on kernel methods require that the bandwidth increases slower than sample size.

However it is well-known that this class of asymptotic estimators does not go well in finite samples, in particular for processes with strong dependence and temporal heterogeneity (see e.g. Müller (2007)). As discussed by Müller (2007), statistical inference by using consistent LRCM estimators performs badly in small samples and also in the presence of dependence and contamination/measurement errors. To overcome these problems, Kiefer and Vogelsang (2002, 2005) introduce a class of kernel-type estimators with a bandwidth rule given by a fixed portion of the sample size, known as fixed- b estimators. These LRCM estimators are not consistent due to a fixed bandwidth parameter. However, they explicitly incorporate parameter uncertainty and they present very good finite sample properties in hypothesis testing. See Kiefer and Vogelsang (2002, 2005) for details.

In particular, Kiefer and Vogelsang (2002) suggest the use of the whole sample as a possible bandwidth rule in the construction of the LRCM estimator (3.1). This strategy allows us to use all lags of X in (1.6) in the estimation procedure. This will be particularly important for noisy data sets with contaminations induced by market microstructure effects and interpolation procedures **(A-B)** discussed in the Introduction.

Let \hat{u}_t be the mean adjusted deviations of the series w_t . The Kiefer and Vogelsang (2002) estimator is defined by

$$\hat{V}_{lr}^{VK} := T^{-1} \sum_{i=1}^T \sum_{j=1}^T \hat{u}_i \left(1 - \frac{|i-j|}{T} \right) \hat{u}_j, \quad (3.2)$$

which corresponds to the use of a Bartlett kernel and bandwidth equal to the sample size in the LRCM estimation.

In order to handle dependent error structures and outliers, Müller (2007) extends the results in Kiefer and Vogelsang (2005, 2002). He shows that the usual LRCM estimators are extremely fragile in the presence of contamination and outliers. He suggests a class of LRCM estimators with bandwidths based on fixed portions of the sample size. The class of estimators proposed by Müller (2007) can be constructed in the same spirit of Kiefer and Vogelsang (2002, 2005) but with one fundamental difference: They are asymptotically robust to contaminations in the autocorrelation structure, in particular to contamination by moving average process. In particular, he obtains a class of estimators $\hat{V}_{lr}^{UA(p)}$ (see pages 1339-1342 in Müller (2007)) which trades optimally robustness and efficiency. The $\hat{V}_{lr}^{UA(p)}$ estimator is defined by

$$\hat{V}_{lr}^{UA(p)} := \frac{\sum_{i=1}^T \sum_{j=1}^T \hat{\xi}_i \hat{\xi}_j}{p}, \quad (3.3)$$

where $\hat{\xi}_t$ is the residual of the linear regression of \hat{u}_t against a p -dimensional series $\hat{v}_t(l)$, $l = 1, \dots, p$ where $\hat{v}_t(l) := \sqrt{2/T} \cos(l\pi(t-1/2)/T)$. The parameter $1 \leq p < \infty$ controls the bias and efficiency in the LRCM estimation procedure. Müller (2007) shows that this estimator has good properties w.r.t robustness and efficiency. In addition, they present well-known asymptotic distributions under the null in hypothesis testing.

In this article, we show that the LRCM estimators (3.2) and (3.3) allow us to filter the dependence structure generated by observational errors and market microstructure effects, typically founded in forward rate markets.

4. Simulation study on the number of factors

In order to investigate the impact of observational errors in the classical PCA methodology applied to forward rate curves, we now provide a detailed simulation study as follows. In order to

study different types of data contamination in forward rate curves, we assume that both the yield and forward rate curves are subject to errors

$$X_t = r_t + \varepsilon_t, \quad Z_t = y_t + \eta_t; \quad t = 0, \dots, T, \quad (4.1)$$

where ε_t and η_t are the error components to be specified.

In addition to the analysis of the curves (4.1), it is also important to consider the first-difference of the time series $\tilde{X}_t := X_t - X_{t-1}$ and $\tilde{Z}_t := Z_t - Z_{t-1}; t = 1, \dots, T$. We recall that a direct application of the PCA to the processes (4.1) implicitly assumes that X and Z are already weakly stationary. If X and (or) Z are assumed to be non-stationary, it is necessary to apply the PCA decomposition at the first-differences \tilde{X} and \tilde{Z} , which are possibly weakly stationary processes. There is also an intrinsic reason to study the first-difference of the time series. Due to the HJM representation of the forward rate (see (1.2)), the PCA methodology must be computed in terms of the increments rather than the first levels. See Jarrow (2002) and Filipovic (2009) for further details.

An obvious consequence of Proposition 2.1 is the following corollary.

Corollary 4.1. *Assume that both r and y are independent weakly stationary processes. If $\eta = \varepsilon = 0$, then the number of principal components of X and Z must be identical.*

The number of principal components will be the key parameter of study in order to infer the performance of the usual PCA methodology based on \hat{V}_s against the LRCM estimators defined in (3.1), (3.2) and (3.3). Of course, the same statement of Corollary 4.1 holds for the first-difference processes \tilde{X} and \tilde{Y} .

In order to investigate the impact of observational errors in the PCA methodology, we proceed the analysis on two different prominent models: Gaussian HJM and Cox-Ingersoll-Ross models. As far as the Gaussian HJM model r_t (see (1.4)) is concerned, the volatility parameter is calibrated based on the classical interest rate curves studied by Diebold and Li (2006). It consists of zero coupon (Treasury bond) bonds with maturities 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 90, 108 and 120 months covering 1985-Jan to 2000-Dec. The Gaussian HJM simulation is based on this calibrated volatility parameter where the number of principal components is equal to 3 (three).

The specification of the Cox-Ingersoll-Ross model (henceforth abbreviated by CIR model) is based on Chen and Scott (2003). In this case, a three-factor CIR model is simulated by means of a short rate process of the form $short_t = \sum_{i=1}^3 Y_t^i$ where each latent factor follows the stochastic differential equation

$$dY_t^i = \kappa_i(\theta_i - Y_t^i)dt + \sigma_i\sqrt{Y_t^i}dB_t^i; \quad i = 1, 2, 3.$$

The parameters κ_i, θ_i and σ_i used in the simulation study are chosen according to Chen and Scott

(2003) who estimate them based on weakly data (1980-1988) of the U.S Treasury market¹. We perform 1,000 replications of these HJM and CIR data generating processes with sample size 1,000.

We analyze the PCA methodology based on the standard sample covariance matrix estimator \hat{V}_s against the LRCM estimators \hat{V}_{lr}^A , \hat{V}_{lr}^{VK} and $\hat{V}_{lr}^{UA(p)}$. Those estimators are applied to (X, Z) and (\tilde{X}, \tilde{Z}) . The \hat{V}_{lr}^A estimator is specified with the quadratic-spectral kernel function and the optimal bandwidth choice of Andrews (1991). The \hat{V}_{lr}^{VK} estimator is specified with the Bartlett kernel function and the whole sample as a bandwidth rule. The $\hat{V}_{lr}^{UA(p)}$ estimator is specified with $p = 4$ components in the basis \hat{v}_t . See Section 3 for details.

Figures 1-5 report the mean value of the cumulative R-Squared obtained in the Monte Carlo experiments with the PCA decompositions based on the static covariance matrix and the LRCM estimators. Figure 1 reports the PCA methodology applied to $X = r + \varepsilon$ where $\varepsilon = 0$ and the associated yield curve is computed via (1.1). Figure 2 reports the PCA methodology applied to $Z = y + \eta$ where $\eta = 0$ and the associated forward rate curve is computed via $y_t(x) + x \frac{\partial y_t(x)}{\partial x}$. Without the presence of observational errors, we notice that the number of principal components is correctly estimated for (X, Z) and (\tilde{X}, \tilde{Z}) by using any of the estimators. This result is not surprising due to Proposition 2.1 and the fact that both r and y are weakly stationary processes. In the presence of observational errors the picture is rather different.

4.1. PCA with measurement error

We formulate the analysis based on two types of observational errors: Market microstructure effects (MME) and interpolation error structure (IES) as described by **(A-B)** in the Introduction. In order to analyze the impact of MME in the PCA methodology, we introduce an additive error structure typically founded in interest rate markets due to transactions costs and liquidity premiums in bond prices (see e.g Mizrach and Neely (2011) and Goyenko et al. (2011)). This phenomenon is directly observed by the existence of bid and ask prices. In the classical model of market microstructure (e.g. Hasbrouck (1991)), the true price of the asset is within the range between the observed bid and ask prices. In other words, the observed prices may be considered as the true price plus an additive measurement error.

We are going to study the impact of the MME as follows. In the sequel, the forward rate process r in (4.2) and the yield process y in (4.3) are given, respectively, by the Gaussian HJM and CIR models as specified above. Figure 3 reports the PCA methodology applied to

$$\begin{cases} X_t(x) = r_t(x) + \varepsilon_t(x); & \text{where } \varepsilon \text{ is a Gaussian zero mean IID process with variance } .0035 \\ Z_t(x) = y_t(x) + \eta_t(x); & \text{where } y \text{ is computed via formula (2.1) and } \eta_t(x) = \frac{1}{x} \int_0^x \varepsilon_t(r) dr. \end{cases} \quad (4.2)$$

¹See Chen and Scott (2003) for the expressions for the affine factors and bond prices in the multifactor CIR model.

Figure 4 reports the PCA methodology applied to

$$\begin{cases} Z_t(x) = y_t(x) + \eta_t(x); & \text{where } \eta \text{ is a Gaussian zero mean IID process with variance } .0035 \\ X_t(x) = r_t(x) + \varepsilon_t(x); & \text{where } \varepsilon_t(x) = \eta_t(x) + x \frac{\partial \eta_t}{\partial x}(x), \quad r_t(x) = y_t(x) + x \frac{\partial y_t}{\partial x}(x). \end{cases} \quad (4.3)$$

We remarkably notice (see Figures 3 and 4) the same structure reported by the empirical analysis given in Liu (2010) when using the classical static covariance estimator \hat{V}_s in the PCA methodology. We observe a substantial difference between the number of factors capable to explain the covariance structure of the yield and forward rate curves, specially when applied to the first-difference of the curves.

In Figures 3 and 4, the estimates based on the standard sample covariance matrix \hat{V}_s are highly biased for the forward rate curves. This can also be observed in the first-difference of the yield curves. We remarkable notice that both \hat{V}_{lr}^{VK} and $\hat{V}_{lr}^{UA(p)}$ estimate the correct number of factors, especially when computed on the first-difference of the forward rate curves in which the contamination is more problematic. The \hat{V}_{lr}^A estimator still indicate an excessive number of factors for the forward rate curves.

Figure 4 reports one typical situation of MME founded in practice. One should notice that the presence of additive IID Gaussian measurement errors in the observed yield curve does not affect the PCA estimation. Neither the usual sample static covariance matrix \hat{V}_s nor \hat{V}_{lr}^A , \hat{V}_{lr}^{VK} and $\hat{V}_{lr}^{UA(p)}$ are affected by the presence of this type of error in the yield curve. In contrast, the estimates based on \hat{V}_s for the forward rate and first-difference of the yield curves are clearly biased. The situation is even worst when dealing with forward rate first-differences.

The explanation for the results reported in Figures 3 and 4 lies in the following argument. We readily see the existence of MME in bond markets impacts differently forward rate and yield curves. In (4.3), the MME introduces a moving average structure with negative persistence into the forward rate curve due to the components $\frac{\partial y_t}{\partial x}$ and $\frac{\partial \eta_t}{\partial x}(x)$. One should notice that the existence of a moving average error structure affects severally the usual estimators of the covariance matrix. See Müller (2007) and Vogelsang and Wagner (2013) for a discussion about this issue. In contrast, the MME in (4.2) only introduces an IID component into the observed yield curve process. This IID component does not affect the temporal dependence of the yield curve. The results reported in Figures 1-4 show that the PCA methodology applied to forward rate curves is a very delicate issue because it might be subject to non-negligible observational errors due to MME. The presence of MME explains the large difference between the number of components indicated by the PCA method for yield and forward curves. This remark is particularly more evident when double differentiation of the underlying moving average process appears due to the first-difference of the observed forward rate curves in (4.3). The main reason to use LRCM estimators instead of the usual \hat{V}_s is to handle

these types of observational errors. In particular, our results strongly indicate the estimators \hat{V}_{lr}^K and $\hat{V}_{lr}^{UA(p)}$ have successfully filtered the MME in the PCA methodology.

Let us now treat the case of a non-additive error structure $G(y_t, \eta_t)$ due to interpolation of the yield curve y_t as described in **(A-B)**. Figure 5 reports the PCA methodology applied to

$$\begin{cases} Z_t(x) = G(y_t(x), \eta_t(x)); & \text{where } G \text{ and } \eta \text{ are generated by an IES induced by a cubic splines} \\ X_t(x) = Z_t(x) + x \frac{\partial Z_t}{\partial x}(x), \end{cases} \quad (4.4)$$

where y is given by the CIR model as specified above. In this analysis, the error is generated by the classical spline method of McCulloch (1975). We simulate bond prices generated by the CIR model and we omit prices for 4 maturities randomly chosen, generating again 1,000 replications for this experiment. The omitted prices are then interpolated by cubic spline using the price curve, similar to the method of McCulloch (1975). Hence, observational errors for bond prices are then generated by means of interpolation. From these prices, we generate yield and forward rate interpolated curves. The PCA method based on the estimators of Section 3 is then applied to $(Z, X, \tilde{Z}, \tilde{X})$ as specified in (4.4).

In Figure 5, the observational errors generated by the cubic splines do not significantly affect the number of factors in yield curves neither in level nor first-difference. However, it has again a nontrivial impact on the forward rate first-difference \tilde{X} . The number of factors of \tilde{X} computed by the PCA based on \hat{V}_s and \hat{V}_{lr}^A is not correctly estimated. The standard PCA method based on \hat{V}_s indicates the incorrect number of six factors which explain 99 % of total variation. Similar to the analysis of MME given in Figures 3 and 4, the PCA based on \hat{V}_{lr}^{VK} and $\hat{V}_{lr}^{UA(p)}$ again correctly estimates the true number of three factors.

The results obtained from these experiments show that the presence of observational errors generated by MME or IES induce a significant bias in the classical PCA methodology applied to forward rate curves. More importantly, our analysis shows that the LRCM estimators $(\hat{V}_{lr}^{VK}, \hat{V}_{lr}^{UA(p)})$ indicate the correct number of factors in the Monte Carlo experiments. The results of this section are robust w.r.t to model specification and types of observational errors.

5. Empirical analysis for the numbers of factors

In this section, we compare the results described in Section 4 with a PCA application to two distinct real data sets. The first one is given by Treasury bonds (zero coupon) with maturities 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 90, 108 and 120 months (17 maturities), with monthly observations ranging from 1985-Jan to 2000-Dez. This data set is constructed based on the Fama-Bliss methodology (unsmoothed Fama-Bliss) and it was already used by Diebold and Li (2006). The second data set is the UK term-structure obtained from the Bank of England with maturities

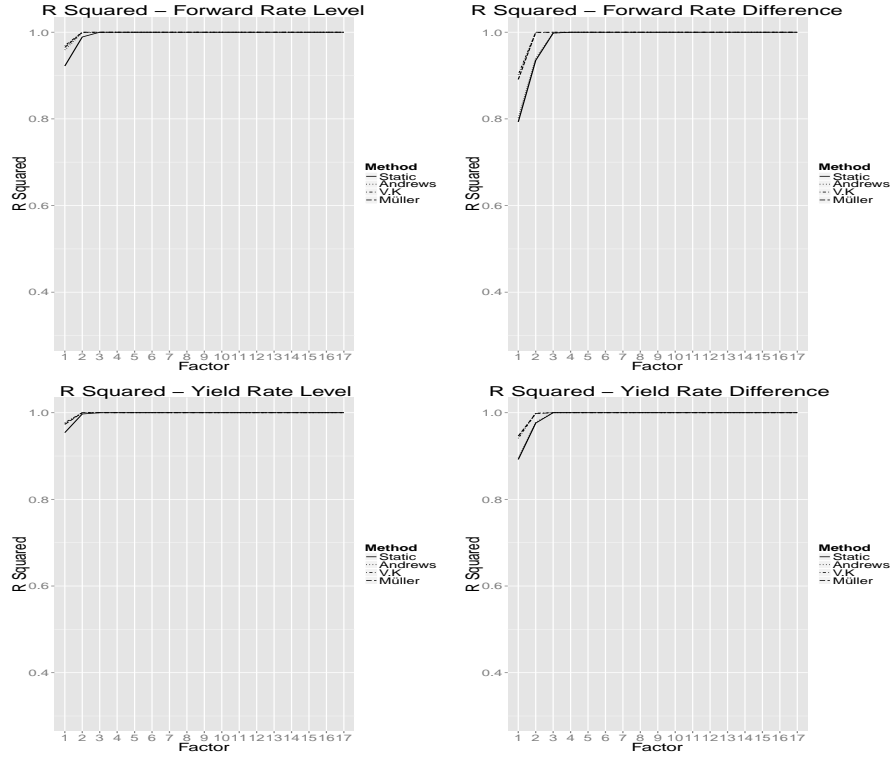


Figure 1: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures. First experiment without observational errors, Gaussian HJM process. Mean values from 1,000 Monte Carlo simulations.

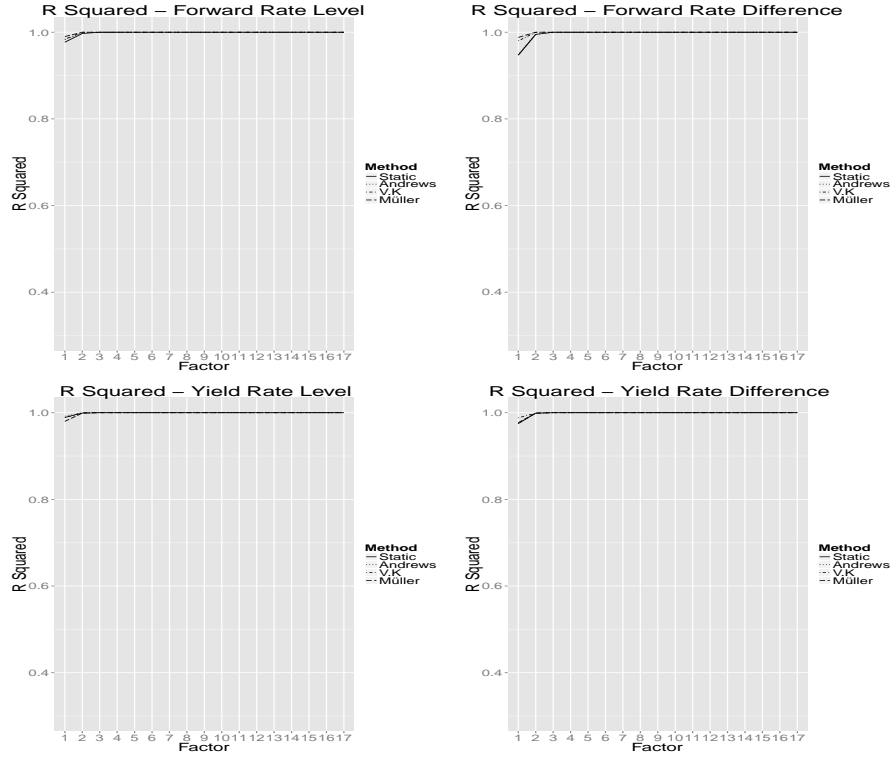


Figure 2: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures . Second experiment without observational errors, Cox-Ingersoll-Ross process. Mean values from 1,000 Monte Carlo simulations.

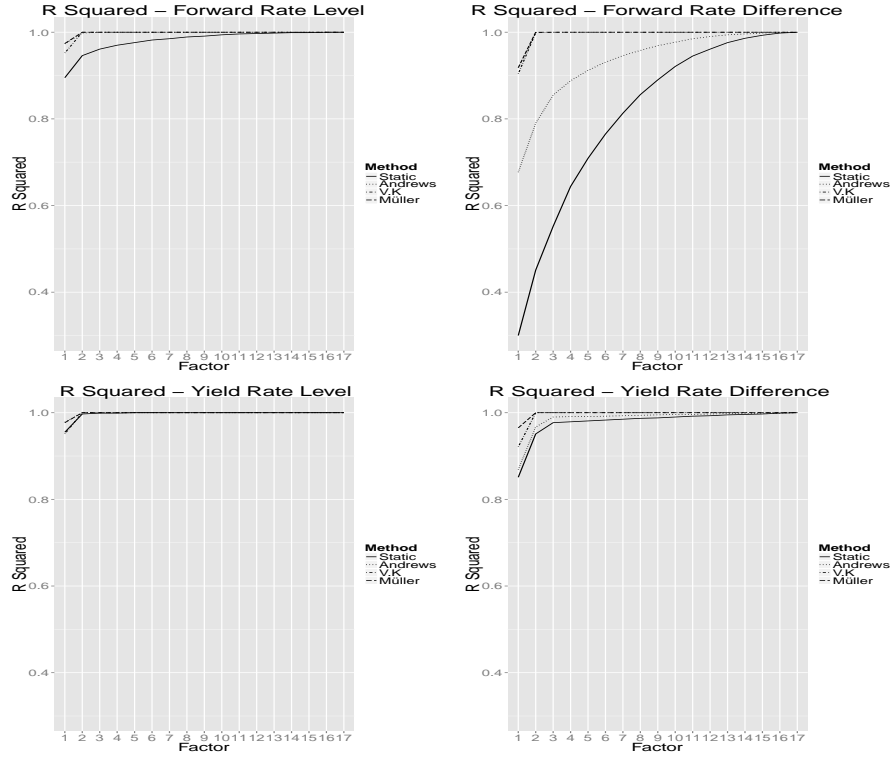


Figure 3: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures. First experiment with observational errors generated by market microstructure effects, Gaussian HJM process. Mean values from 1,000 Monte Carlo simulations.

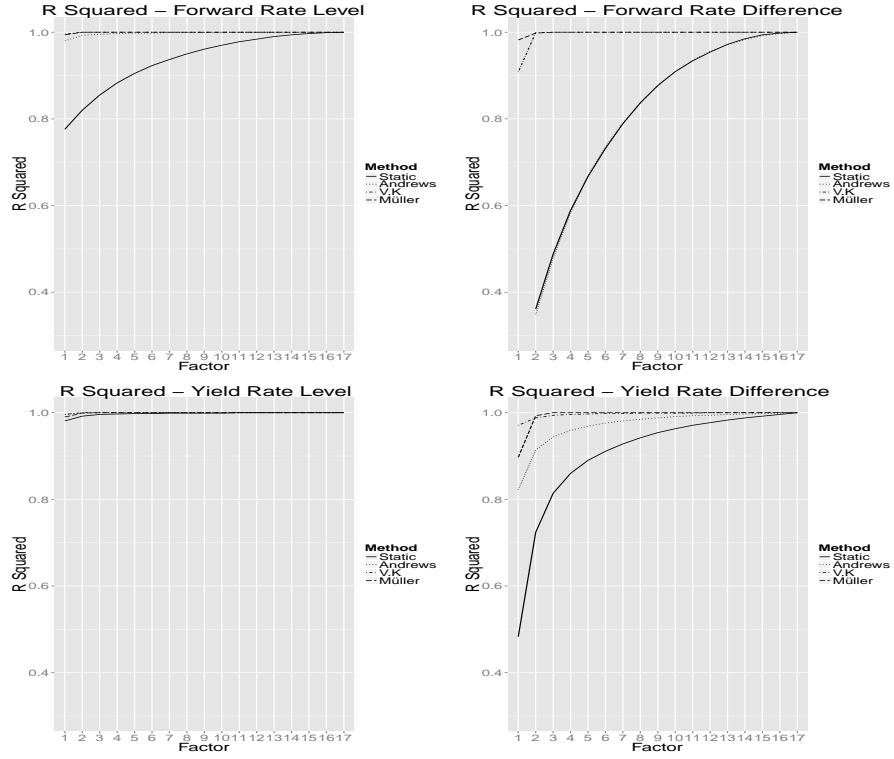


Figure 4: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures. Second experiment with observational errors generated by market microstructure effects, Cox-Ingersoll-Ross process. Mean values from 1,000 Monte Carlo simulations.

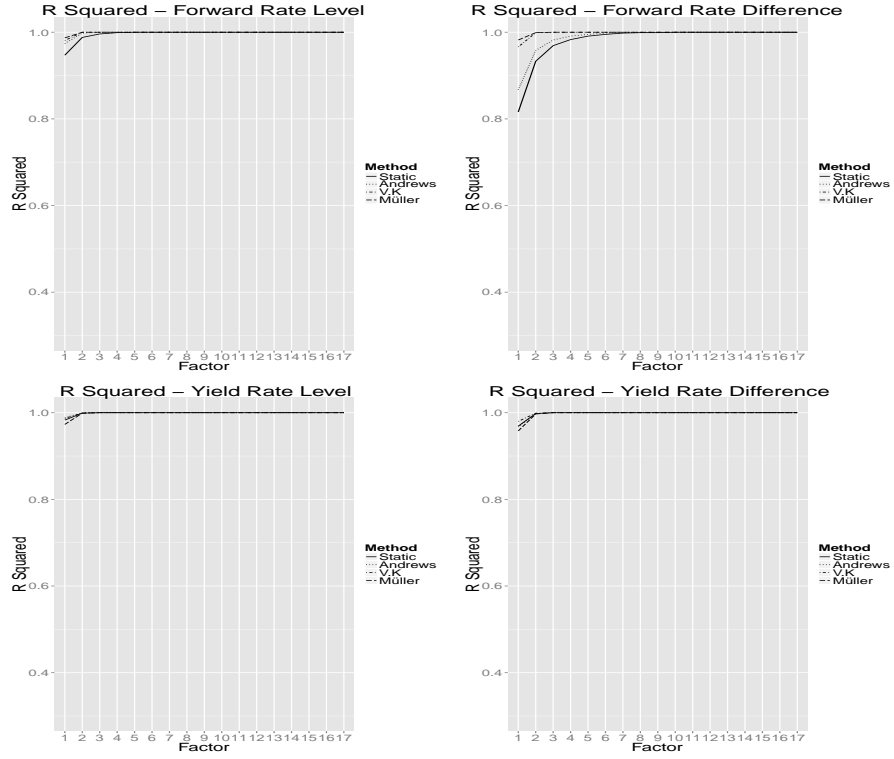


Figure 5: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures. Experiment with observational errors generated by cubic spline interpolation, Cox-Ingersoll-Ross process. Mean values from 1,000 Monte Carlo simulations.

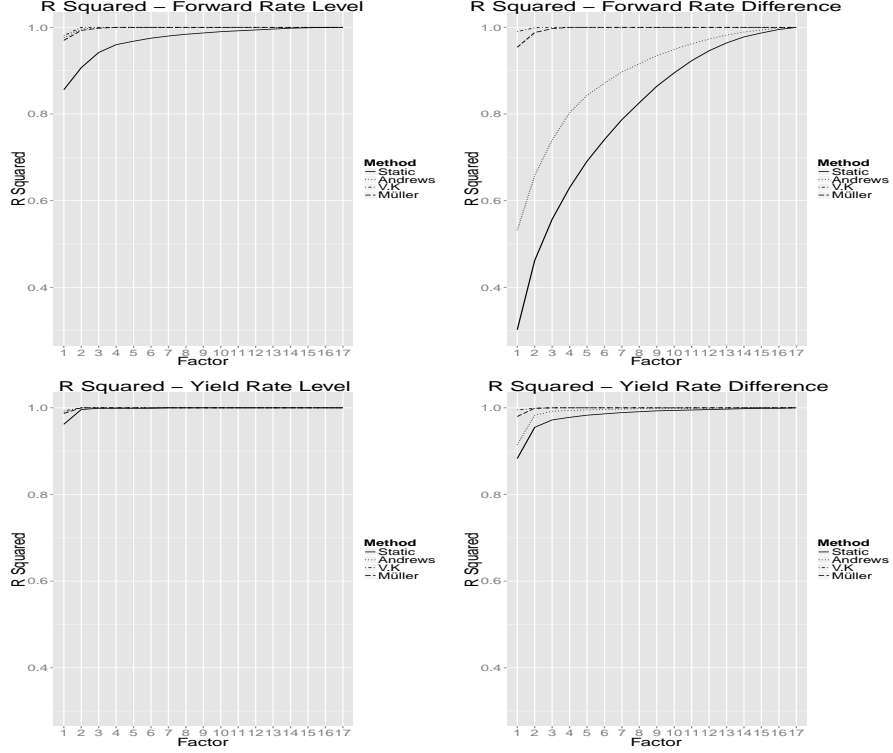


Figure 6: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures - U.S. Treasury Curve - Fama-Bliss Database.

.5 to 25 years (50 maturities) and daily data ranging from Jan/4/2005 to Feb/29/2012 summing 1872 observations.

In Figures 6 and 7, we report the cumulative R^2 obtained from the PCA decomposition for the forward rate and yield curves of the U. S. and U. K. markets, respectively. The usual static covariance matrix estimator \hat{V}_s applied to these two data sets shows the same behavior as described by Liu (2010). We observe a large difference for the number of principal components between yield and forward rate curves, both in the first-difference and level. In particular, we observe the same type of behavior as described in Section 4. More importantly, we notice that both estimators \hat{V}_{lr}^{VK} and $\hat{V}_{lr}^{UA(p)}$ indicate the same number of factors between yield and forward rate curves. This result holds for both the first-difference and the level of the curves. More importantly, it corroborates with our findings in Figure 4.

According to Proposition 2.1, these results strongly suggest the presence of contamination errors in the data. In the US data, the error contamination might be explained by the Fama-Bliss method of construction of the forward rate curve (Fama and Bliss (1987)) which uses a piecewise constant function to approximate the discount factor. The Bank of England term structure is constructed

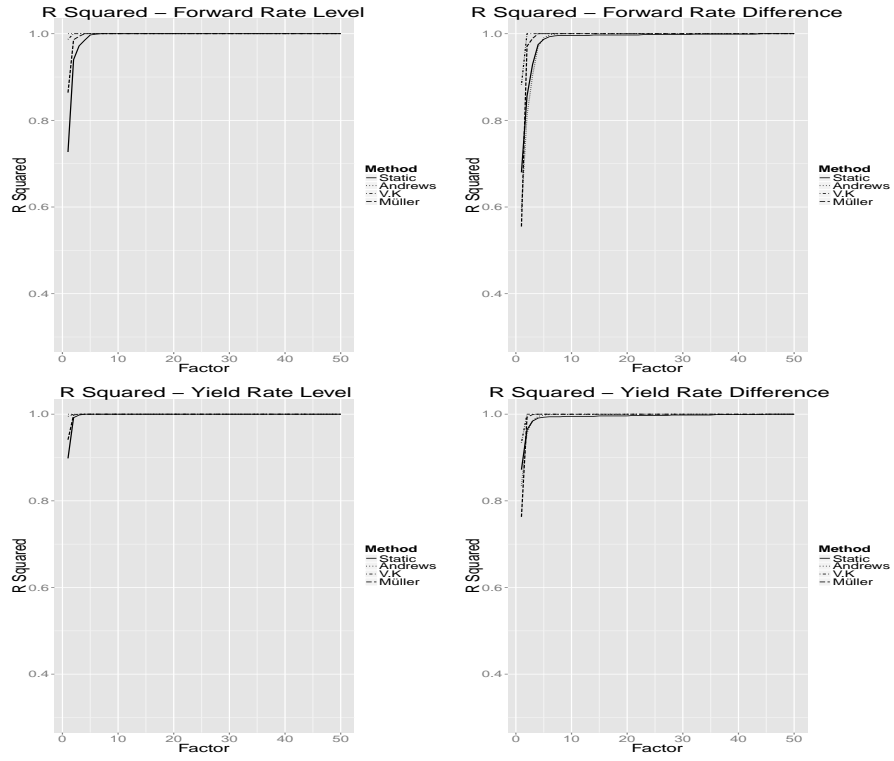


Figure 7: Cumulative R^2 obtained from the PCA decomposition for the level and first difference of the forward and yield term structures - U.K. Bank of England Term Structure Database.

by means of a smooth cubic spline interpolation method for the yield and forward rate curves (see e.g. Anderson and Sleath (2001) for details). Therefore, it is subject to observational errors similar to the Monte Carlo experiment (4.4) in Figure 5. The presence of MME in the US data might also be a source of observational errors as reported by Goyenko et al. (2011) and Mizrach and Neely (2011). The reader is urged to compare the Monte Carlo experiments (4.3) and (4.4) to Figures 6 and 7, respectively. The empirical analysis of this section jointly with Section 4 and Proposition 2.1 give strong support for the use of PCA based on the LRCM estimators described in Section 3.

6. Pricing interest rate derivatives

In Section 4, we study the impact of noisy data in the PCA methodology. In Section 5, we report the existence of nontrivial measurement errors in forward rate markets. In this section, the primary goal is to illustrate the importance of a correct spectral analysis in noisy forward rate curves. The example we choose to illustrate this point is the pricing of interest rate derivatives in the presence of noisy data. We study the impact of observational errors in pricing interest rate derivatives by means of the PCA methodology. More precisely, we compare the pricing error of the PCA methodology based on \hat{V}_s against the LRCM estimators.

Similar to Mercurio and Moraleda (2000), we consider a vanilla-type call option based on a zero coupon bond with 10 years maturity, with time to maturities of .25, .5 and 1 years, and distinct strike prices. The underlying data generating process is specified by a 2-factor Hull-White (g2++) model (see e.g. pages 356-364 in Nawalkha et al. (2007)) which yields an analytic pricing formula for vanilla options written on a zero coupon bond. This model is affine whose the correspondent forward rate volatility function is given by

$$Vol[df(t, T)] = \sqrt{v_1^2 e^{-\kappa_1(T-t)} + v_2^2 e^{-\kappa_2(T-t)} + 2v_1 v_2 \rho_{12} e^{-(\kappa_1 + \kappa_2)(T-t)}}; \quad 0 \leq t \leq T,$$

where the parameters $v_1, v_2, \kappa_1, \kappa_2, \rho_{12} \in \mathbb{R}$. The payoff of interest is given by

$$V_{T^0} := \max \{P(T^0, T) - K, 0\}$$

for some $T^0 < T$, where T is the time of exercising the payoff V_{T^0} . The Gaussian structure of the 2-factor Hull-White model yields a closed form expression for the price $\mathbb{C}^0(T^0, T)$ of V_{T^0} as follows

$$\mathbb{C}^0(T^0, T) = P(0, T)\Phi(d_+) - P(0, T^0)\Phi(d_-),$$

where Φ is the Gaussian cumulative distribution function, $d_{\pm} := \frac{\ln \frac{P(0, T)}{K P(0, T^0)} \pm \frac{v}{2}}{\sqrt{v}}$ and v is the integrated volatility between $[T^0, T]$ which follows equation (4.41) in Andersen and Piterbarg (2010), page (185).

We construct the interest rate curves with maturities 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 40, 60, 72, 90, 108 and 120 months with delta of order 1/252 (daily data) between each curve observation. In order to verify the robustness of our results, we perform a Monte Carlo study with two distinct set of parameters - the first one consist $\kappa_1 = .8, \kappa_2 = .7, v_1 = .1, v_2 = .1$ and $\rho_{12} = -.3$, and the second one $\kappa_1 = .9, \kappa_2 = .85, v_1 = .1, v_2 = .2$ and $\rho_{12} = -.3$. The Monte Carlo experiments are based on the observational process $Z = y + \eta$, where y is the correspondent yield curve process specified by the above 2-factor Hull-White model and η is zero or an IID zero mean Gaussian noise with variance .0035. The error is understood in terms of Mean Square Error (MSE)

$$MSE (Price_{pca(Z)}(T^0, T), \mathbb{C}^0(T^0, T))$$

where $Price_{pca(Z)}(T^0, T)$ is the price computed by means of the estimation of the integrated volatility v based on PCA as a function of the sample covariance matrix \hat{V}_s and the LRCM $\hat{V}_{lr}^A, \hat{V}_{lr}^{VK}$ and $\hat{V}_{lr}^{UA(p)}$, given, respectively, by formulas (3.1), (3.2) and (3.3).

Table 1 reports the Mean Square Error pricing with the first set of parameters, $\eta = 0$ and based on a sample size of 250 observations. The PCA pricing using the sample covariance matrix \hat{V}_s and the estimators \hat{V}_{lr}^A and $\hat{V}_{lr}^{UA(p)}$ presents similar results in terms of pricing errors, while the \hat{V}_{lr}^{VK} estimator achieves the best results. This can be explained by the excellent finite sample properties of \hat{V}_{lr}^{VK} compared with the other estimators. See Kiefer and Vogelsang (2005, 2002) for further details.

The results reported in Tables 2 and 3 suggest that the pricing of interest rate derivatives based on PCA with sample covariance matrix \hat{V}_s is sensitive to observational errors. Clearly, the pricing based on the LRCM estimators $\hat{V}_{lr}^A, \hat{V}_{lr}^{VK}$ and $\hat{V}_{lr}^{UA(p)}$ presents better performance. We clearly observe smaller Mean Squared Errors for all strikes and time to maturities. This restates the validity and relevance of the corrections w.r.t the estimation of the covariance matrix in the PCA methodology.

The estimator \hat{V}_{lr}^{VK} presents the best performance among the methods discussed in Section 3. This result is particularly important because it is the simplest one to implement where no bandwidth parameter estimation is needed. These results strongly suggest that neglecting the underlying dependence structure generated by observational errors may degenerate significantly the pricing of interest rate derivatives. We also present in Table 4 the experiment with the first parameter set and with contaminations errors, but now using a sample of 1,000 observations. The results are similar to Tables 2 and 3. The Monte Carlo experiment of this section is robust w.r.t sample size and parameters.

.25 Years to Mat.	Strike	0.45	0.5	0.55	0.6
	Static \hat{V}_s	0.034	0.047	0.053	0.051
	Andrews \hat{V}_{lr}^A	0.033	0.046	0.053	0.050
	V-K \hat{V}_{lr}^{VK}	0.011	0.017	0.020	0.019
	Müller $\hat{V}_{lr}^{UA(p)}$	0.033	0.046	0.052	0.050
.5 Years to Mat.	Strike	0.48	0.53	0.58	0.63
	Static \hat{V}_s	0.012	0.022	0.031	0.033
	Andrews \hat{V}_{lr}^A	0.012	0.022	0.031	0.033
	V-K \hat{V}_{lr}^{VK}	0.004	0.007	0.010	0.011
	Müller $\hat{V}_{lr}^{UA(p)}$	0.012	0.022	0.031	0.032
1 Year to Mat.	Strike	0.5	0.55	0.6	0.65
	Static \hat{V}_s	0.005	0.013	0.021	0.024
	Andrews \hat{V}_{lr}^A	0.005	0.013	0.021	0.024
	V-K \hat{V}_{lr}^{VK}	0.002	0.004	0.007	0.008
	Müller $\hat{V}_{lr}^{UA(p)}$	0.006	0.013	0.021	0.024

Table 1: MSE - Option Pricing on a Zero Coupon Bond - without measurement errors, first parameter set, sample size 250.

7. Final remarks

In this article, we discuss the impact of the inherent presence of measurement errors in the classical application of the PCA methodology in the estimation of forward rate curves. Our results strongly suggest the classical PCA method based on the standard sample covariance matrix is *not* suitable for estimating factors of forward rate curves. The main reason is the appearance of non-negligible observational errors in the implied forward rate curves. An alternative methodology based on so-called long-run covariance matrix estimators seems to improve significantly the quality of the estimation of the principal components of forward rate curves. We illustrate the importance of a correct spectral analysis in forward rate curves by presenting non-trivial effects of noisy data in pricing errors related to European call options. Lastly, the results of this paper yield a sound explanation for the remarkable difference between the estimated number of factors in yield and forward rate curves reported in the literature (see e.g. Akahori et al. (2006), Akahori and Liu (2011), Lekkos (2000) and Liu (2010)).

Our conclusion is supported by the following results presented in this paper. Proposition 2.1 proves that the number of principal components of the observed forward rate and yield curves must be identical in the absence of measurement errors. Based on this fact, we perform a detailed simulation analysis in three prominent interest rate models with a number of distinct parameters in the presence of observational errors of various magnitudes and forms. Market microstructure effects and interpolation errors due to extraction of forward rates from yield curves are carefully analyzed. The results clearly indicate a considerable bias in the classical PCA methodology applied to forward

.25 Years to Mat.	Strike	0.45	0.5	0.55	0.6
	Static \hat{V}_s	0.056	0.069	0.076	0.075
	Andrews \hat{V}_{lr}^A	0.039	0.052	0.058	0.057
	V-K \hat{V}_{lr}^{VK}	0.013	0.019	0.022	0.021
	Müller $\hat{V}_{lr}^{UA(p)}$	0.034	0.046	0.051	0.049
.5 Years to Mat.	Strike	0.48	0.53	0.58	0.63
	Static \hat{V}_s	0.037	0.049	0.059	0.062
	Andrews \hat{V}_{lr}^A	0.020	0.030	0.039	0.041
	V-K \hat{V}_{lr}^{VK}	0.004	0.008	0.011	0.012
	Müller $\hat{V}_{lr}^{UA(p)}$	0.013	0.022	0.030	0.032
1 Year to Mat.	Strike	0.5	0.55	0.6	0.65
	Static \hat{V}_s	0.029	0.041	0.051	0.056
	Andrews \hat{V}_{lr}^A	0.013	0.021	0.030	0.034
	V-K \hat{V}_{lr}^{VK}	0.002	0.004	0.007	0.009
	Müller $\hat{V}_{lr}^{UA(p)}$	0.006	0.013	0.021	0.024

Table 2: MSE - Option Pricing on a Zero Coupon Bond - with measurement errors, first parameter set, sample size 250

rate curves. In contrast, Monte Carlo experiments jointly with empirical analysis strongly suggest that PCA based on long run covariance matrices is robust w.r.t measurement errors in forward rate curves. The presence of observational errors in forward rate markets is validated by Proposition 2.1 together with a detailed empirical analysis on the number of principal components in the UK and US interest rate markets. Monte Carlo experiments reports that classical PCA based on sample covariance matrices presents nontrivial pricing errors for European call options. In contrast, the use of long run covariance matrices seems to correct this bias.

References

- Akahori, J., Aoki, H., Nagata, Y., 2006. Generalizations of Ho-Lee's binomial interest rate model I: from one- to multi-factor. *Asia-Pacific Financial Markets* 13, 151–179.
- Akahori, J., Liu, N.-L., 2011. On a type-I error of a random walk hypothesis of interest rates. *International Journal of Innovative Computing, Information and Control* 7, 115–131.
- Alexander, C., Lvov, D., 2003. Statistical properties of forward LIBOR rates, working paper.
- Andersen, L. F. B., Piterbarg, V. V., 2010. *Interest Rate Modeling, Volume I: Foundations and Vanilla Models*. Atlantic Financial Press.
- Anderson, N., Sleath, J., 2001. New estimates of the uk real and nominal yield curves. The Bank of England Working Paper Series 126, 1–41.

.25 Years to Mat.	Strike	0.45	0.5	0.55	0.6
	Static \hat{V}_s	0.078	0.087	0.091	0.088
	Andrews \hat{V}_{lr}^A	0.066	0.076	0.078	0.075
	V-K \hat{V}_{lr}^{VK}	0.024	0.029	0.030	0.029
	Müller $\hat{V}_{lr}^{UA(p)}$	0.060	0.069	0.071	0.068
.5 Years to Mat.	Strike	0.48	0.53	0.58	0.63
	Static \hat{V}_s	0.048	0.058	0.065	0.065
	Andrews \hat{V}_{lr}^A	0.035	0.044	0.050	0.051
	V-K \hat{V}_{lr}^{VK}	0.009	0.013	0.016	0.016
	Müller $\hat{V}_{lr}^{UA(p)}$	0.027	0.036	0.041	0.041
1 Year to Mat.	Strike	0.5	0.55	0.6	0.65
	Static \hat{V}_s	0.036	0.046	0.054	0.056
	Andrews \hat{V}_{lr}^A	0.022	0.031	0.037	0.039
	V-K \hat{V}_{lr}^{VK}	0.005	0.008	0.011	0.012
	Müller $\hat{V}_{lr}^{UA(p)}$	0.015	0.022	0.027	0.029

Table 3: MSE - Option Pricing on a Zero Coupon Bond - with measurement errors, second parameter set, sample size 250.

- Andrews, D., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.
- Barzanti, L., Corradi, C., 1998. A note on interest rate term structure estimation using tension splines. *Insurance: Mathematics and Economics* 22, 139–143.
- Bhar, R., Chiarella, C., Tô, T.-D., 2002. A maximum likelihood approach to estimation of Heath-Jarrow-Morton models, working paper.
- Chen, R., Scott, L., 2003. Multi-factor Cox-Ingersoll-Ross models of the term structure: Estimates and tests from a Kalman filter model. *Journal of Real Estate Finance and Economics* 2, 143–172.
- Chiu, N.-C., Fang, S.-C., Lavery, J. E., Lin, J.-Y., Wang, Y., 2008. Approximating term structure of interest rates using cubic L1 splines. *European Journal of Operational Research* 184, 990–1004.
- Diebold, F. X., Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 337–364.
- Falini, J., 2010. Pricing caps with HJM models: The benefits of humped volatility. *European Journal of Operational Research* 207, 1358–1367.
- Fama, B., Bliss, R. R., 1987. The information in long-maturity forward rates. *The American Economic Review* 77, 680–692.

.25 Years to Mat.	Strike	0.45	0.5	0.55	0.6
	Static \hat{V}_s	0.059	0.071	0.076	0.073
	Andrews \hat{V}_{lr}^A	0.039	0.050	0.054	0.051
	V-K \hat{V}_{lr}^{VK}	0.009	0.014	0.016	0.016
	Müller $\hat{V}_{lr}^{UA(p)}$	0.033	0.044	0.047	0.044
.5 Years to Mat.	Strike	0.48	0.53	0.58	0.63
	Static \hat{V}_s	0.039	0.051	0.060	0.062
	Andrews \hat{V}_{lr}^A	0.018	0.028	0.035	0.036
	V-K \hat{V}_{lr}^{VK}	0.002	0.005	0.008	0.009
	Müller $\hat{V}_{lr}^{UA(p)}$	0.013	0.022	0.028	0.029
1 Year to Mat.	Strike	0.5	0.55	0.6	0.65
	Static \hat{V}_s	0.031	0.043	0.052	0.056
	Andrews \hat{V}_{lr}^A	0.011	0.019	0.027	0.029
	V-K \hat{V}_{lr}^{VK}	0.001	0.003	0.005	0.006
	Müller $\hat{V}_{lr}^{UA(p)}$	0.006	0.013	0.019	0.021

Table 4: MSE - Option Pricing on a Zero Coupon Bond - with measurement error, first parameter set, sample size 1,000

Filipovic, D., 2009. Term-Structure Models: A Graduate Course. Springer Finance.

Gauthier, G., Simonato, J.-G., 2012. Linearized Nelson-Siegel and Svensson models for the estimation of spot interest rates. European Journal of Operational Research 219, 442–451.

Goyenko, R., Subrahmanyam, A., Ukhov, A., 3 2011. The term structure of bond market liquidity and its implications for expected bond returns. Journal of Financial and Quantitative Analysis 46, 111–139.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica 60, 77–105.

Jarrow, R. A., 2002. Modeling Fixed Income Securities and Interest Rate Options. Stanford University Press.

Kiefer, N. M., Vogelsang, T. J. ., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory 21, 1130–1164.

Kiefer, N. M., Vogelsang, T. J., 2002. Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica 70, 2093–2095.

- Lekkos, I., 2000. A critique of factor analysis of interest rates. *Journal of Derivatives* 8, 72–83.
- Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1, 54–61.
- Liu, N., 2010. A comparative study of principal component analysis on term structure of interest rates. *JSIAM Letters* 2, 57–60.
- Lord, R., Pelsser, A., 2007. Level-slope-curvature: Fact or artefact ? *Applied Mathematical Finance* 14, 105–130.
- McCulloch, J. H., 1975. The tax adjusted yield curve. *Journal of Finance* 30, 811–830.
- Mercurio, F., Moraleda, J., 2000. An analytically tractable interest rate model with humped volatility. *European Journal of Operational Research* 120, 205–214.
- Mizrach, B., Neely, C. J., 2011. Treasury market microstructure of the U.S. In: Meyers, R. A. (Ed.), *Complex Systems in Finance and Econometrics*. Springer New York, pp. 863–875.
- Müller, U., 2007. A theory of robust long-run variance estimation. *Journal of Econometrics* 141, 1331–1352.
- Nawalkha, S. K., Believa, A. B., Soto, G. M., 2007. *Dynamic Term Structure Modeling*. John Wiley & Sons, Inc.
- Rebonato, R., 2002. *Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Model and Beyond*. Princeton University Press.
- Rutkowski, M., 1996. Valuation and hedging of contingent claims in the HJM model with deterministic volatilities. *Applied Mathematical Finance* 3, 237–267.
- Schmidt, W. M., 2011. Interest rate term structure modelling. *European Journal of Operational Research* 214, 1–14.
- Stambaugh, R. F., 1988. The information in forward rates: Implications for models of the term structure. *Journal of Financial Economics* 21, 41–70.
- Steeley, J., 1990. Modelling the dynamics of the term structure of interest rates. *The Economic and Social Review* 21, 337–361.
- Vasicek, O., Fong, H. G., 1982. Term structure modeling using exponential splines. *Journal of Finance*, 37, 339–356.
- Vogelsang, T. J., Wagner, M., 2013. A fixed-b perspective on the Phillips-Perron unit root tests. *Econometric Theory* 29, 609–628.